Copyright

by

Norma Estela Palomino

2003

The Bibliographic Concept of *Work* in Cataloguing and its Issues.

by

Norma Estela Palomino

Report

Presented to the Graduate School of Information

Of the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Information Studies

The University of Texas at Austin

May, 2003

The Bibliographic Concept of *Work* in Cataloguing and its Issues.

APPROVED BY

SUPERVISING COMMITTEE:

_____

Francis Miksa, supervisor

_____

Don Gibbs, co-supervisor

_____

Ronald Wyllys, reader

Dedication

to my mother / *a mi madre*

Ofelia Mabel Pisano

who inspired my freedom / *quien me enseñó a ser libre*

to be creative / *para inventar mis mundos*

and

to my father / *a mi padre*

Rodolfo Javier Palomino

who taught me the persistence / *quien me instruyó perseverancia*

to achieve / *para crearlos*

Les agradezco

The Bibliographic Concept of Work in Cataloguing and its Issues.

by

Norma Estela Palomino, MSInfoSt

The University of Texas at Austin, 2003

SUPERVISOR: Francis Miksa

CO-SUPERVISOR: Don Gibbs

Abstract

This report explores the IFLA's document *Functional Requirements for Bibliographic Records* (FRBR). It discusses the notion of *work* in cataloguing as it was built since the 1950s, inasmuch this notion constitutes the conceptual framework for the proposal. Also, the *entity-relationship database modeling* (ERDM) system is described as far as such model provides to FRBR the operative elements that make it functional. ERDM gives to FRBR a user-centered approach as well. In its third chapter, the report tests the FRBR model through its application to a set of items belonging to the novel *Rayuela,* by Julio Cortazar, held at the Benson Latin American Collection of the University of Texas at Austin. Finally, some critical issues are raised along with general conclusions regarding the functionality of the model.

Table of Contents

Introduction

The description of information entities (IEs) often causes cataloguers to make assumptions about the nature of such entities and their relationships. For example, traditional cataloguing has tended to approach each single physical item such as a book, a music score or a map (or each obvious set of physical items such as a two volume encyclopedia or a multiple volume serial) as a unique piece of information—a single IE. Thus, a copy of the book "The old man and the sea" by Ernest Hemingway will be described as a single IE in the library catalogue, without exploring the relationship that the novel has with other copies or editions of that work within the catalogue. When following this procedure, the cataloguer is not ordinarily concerned about whether the library collection holds a translation of the novel into Spanish, an abridged version for students, an edition with comments made by a specialist in Hemingway's writing style and philosophical ideas, and so on.

However, each of these versions of the same work involves different and important information for the reader while choosing which copy of the book he wants to borrow. For a student writing a paper about the novel, for example, an edition with comments made by a specialist might be highly appreciated. Nowadays, the search for such an edition will probably be done in two or three steps, since the catalogue does not always clearly link different editions to each other. Instead, the user has to search the title in the catalogue and, after that,

carefully read the detailed information in the description (e.g., the statement of responsibility, edition, or notes' fields) to see what relationships are involved, if indeed the information is given at all. Similarly, the relationship of augmented or condensed editions to the original text of the novel is not always made clear in the surrogate record.

Actually, the problem of multiple versions of a novel like "The old man and the sea" critically increases in the case of classical books, like the "Iliad" or the "Odyssey," by Homer, which have uncountable editions, versions, critical analyses, etc. A wide range of users (including specialists such as linguists, philologists or philosophers) may want to browse in our catalogs for different editions and visualize those editions in a comprehensive way. For example, a user may want to browse the catalogue for the "Iliad" and retrieve surrogate records ordered by edition numbers or edition types (augmented, condensed, annotated, etc.). A tool for retrieving that provided such information would help the user to find the information needed faster and clearly. Also, a tool of this kind would maximize the *precision* (or the quality of retrieving only relevant records as a result for a search) and minimize the *recall* (or the function of retrieving all records related to the search, even those with a weak relationship) of a search result (Boop and Smith, 2001).

Unfortunately, traditional library cataloguing did not come up with an appropriate approach for describing items that clearly show such diversity of

relationships to the same artistic or intellectual creation. Since the 1950s, however, cataloguing researchers have been exploring a new baseline to solve this problem. Cataloguing IEs has been facing difficult issues for several decades. Standards, rules and formats (such as ISBD, AACR2 and MARC) built the tradition of cataloguing from an approach based on the individual description of single items, disregarding the potential relationships among those single items within the catalog. The relationships among such items become essential in the case of intellectual or artistic creations that constitute fundamental contributions in the history of science, literature, and arts. Throughout history, individuals around the world have generated multiple versions and re-creations of those fundamental works, such as the *Iliad* and the *Odyssey* by Homer, *Hamlet* by Shakespeare, or *Don Quijote de la Mancha* by Cervantes.

Several authors, among whom are Seymour Lubetzky (1969), Elaine Svenonius (2000) and Richard Smiraglia (2001) participated in building a fundamental theory for a new approach to cataloguing. This theory includes the fundamental concepts of *work, superwork* and *bibliographic family of entities*. More particularly, this theory draws an important difference between the *content* or abstract component of a book and the *container* or physical medium in which a *work* is embodied. As we will discuss later, the conceptualization of an abstract level or *content* for a book, as differentiated from the concrete or

physical embodiment in which it is carried, will lead to the recognition of each *work* as related to but essentially independent of its several physical embodiments—the item which has been the traditional starting point for catalog descriptions.

Recently, in 1998, the International Federation of Library Associations (IFLA) identified this trend in cataloguing as a possible solution for the above discussed information-retrieval issues, along with other problems faced by the international library community such as those related to reducing costs of cataloguing through the creation of "less-than-full-level records" (IFLA, 1998, p.2). In the document *Functional Requirements for Bibliographic Records* (FRBR), IFLA argues that the new approach to cataloguing will provide libraries with a briefer surrogate record, reducing the length of surrogate records to the minimum level of description required for users to identify and retrieve an item. IFLA's rationale through the FRBR document is that the new trend in describing IEs will focus on *what the user looks for* in an item rather than on *what can be extensively described* about each physical item. In this way, cataloguers will be helped to avoid making meticulous, but essentially duplicated, surrogate records.

The goal of my report is to introduce, discuss, and evaluate this new approach for describing bibliographic items, centered in the concept of *work* represented by the above-mentioned authors along with the FRBR by IFLA. My

report is arranged in four chapters. In the first chapter I will discuss the basic concepts related to the new approach as they were expressed by authors Seymour Lubetzky, Elaine Svenonius and Richard Smiraglia. These concepts are: *work, superwork,* and *bibliographic families.* The second chapter introduces the document by IFLA, which begins with the concept of *a work* but introduces also the *entity-relationship database model* (ERDM) as an instrumental approach to an operational model for cataloguing (from now on called *the FRBR model*). In chapter three, I will apply the FRBR model to the holdings of a particular work in a library, the novel *Rayuela* by author Julio Cortázar. This section proposes surrogate records for a set of IEs of that novel held by the Benson Latin American Collection of the University of Texas at Austin. A final chapter explores the advantages and limitations of the FRBR model, elaborated through my experience of applying FRBR to the novel *Rayuela* and draws general conclusions and raises challenges for the future development of this fascinating model.

Chapter One: Basic Concepts

Researchers Seymour Lubetzky, Elaine Svenonius and Richard Smiraglia are three important reviewers of practices in cataloguing. These authors attempted to rebuild cataloguing principles by thoroughly studying the nature of IEs along with the assumptions that traditional cataloguing made while describing them. The book by Elaine Svenonius *The intellectual foundation of information organization* (2000) constitutes a landmark work for any researcher interested in the topic. Svenonius explores the notion of *a work* in the context of the foundations of IE organization, but also introduces the concept of *superwork.* Her ideas have been strongly influenced by Lubetzky's approach to the idea of *a work*. Additionally, Richard Smiraglia develops the notion of *bibliographic family* of IEs, a concept that is closely related to *superwork* as we will see. In this chapter, I will discuss these three fundamental notions—*work, superwork* and *bibliographic family of IEs.*

1.1. *Work*.

Throughout the development of the practice of cataloguing, the notion of work has been a controversial one. Lubetzky (1969) differentiated *work* from *book* and clearly stated the nature of the former, *work,* as the focus of the cataloguers' description efforts. According to Lubetzky, *a work* refers to the

intellectual or artistic content of the physical item that we hold in the library. According to this notion, the ideas, creations and meanings elaborated by one or several creators, i.e. the *work*, should be detached from their physical embodiment, which should properly be called *a book*. As Lubetzky says, "the book … comes into being as a dichotomous product—as a material object or medium used to convey the intellectual work of an author. Because the material book embodies and represents the intellectual work, the two have come to be confused" (Lubetzky, 1969, p. 11). As we will discuss, the distinction between these two components in every item (i.e. the abstract content distinguished from the concrete container) constitutes a landmark in changing the traditional approach to cataloging.

Elaine Svenonius takes this basic differentiation (between *work,* or the *content* of an IE, and its physical item or *container*) as the baseline for building an epistemological and ontological theory for the classification and description of IEs. Taking this notion of the abstract component, Svenonius draws a broad definition for the concept of *work.*

Svenonius defines *a work* formally as "the set of all documents that are copies of (equivalent to) a particular document a_i (an individual document chosen as emblematic of the work, normally its first instance) or related to this individual by revision, update, abridgment, enlargement, or translation" (Svenonius, 2000, p.37). Following this notion, there are two kinds of items that

belong to the *work* set: (i) items that are equal to the original *work,* and (ii) those items that are related but are not exactly equivalent to the original *work.* The two sets of items are differentiated by their level of similarity to the original work, the former set (the *copies* of the work) having a closer relationship while the latter set is wider and does not require the items to be exactly equivalent to the parent work. Therefore, Svenonius' broad definition includes a wide range of items in the total set encompassed by each work. Her notion of a work refers (as in Lubetzky) to the abstract content, giving a common origin to the set of all items that are essentially the same, but at the same time, keeping them together and related among each other.

At this point, an example can help to depict the two set of items related to a *work* (the set of items *equivalent to* and the set of items *essentially the same as* the parent work). We will find in a library catalog the work "The heart is a lonely hunter" (by Carson McCullers), conveyed by a wide set of printings made by many publishers. For example, this novel was published in New York by Modern Library publishers and in Boston by Houghton Mifflin Company. The former printing has 430 pages, while the latter has 356. However, beyond physical variances such as number of pages, binding quality (paperback or hard cover), font size or type, etc., the text of the content, i.e. Carsons McCullers' original creation, remains the same. This content is also the shared characteristic that gathers both printing copies together in reference to a common origin. In

consequence, we have here two examples of items within the work set that are *equivalent* to the original work in the sense that they convey the same information.

We might also find in the same catalogue other items that are also manifestations of this work, for example, a translation into German entitled *Das Herz ist ein einsamer Jager, roman,* and perhaps also an abridged edition, and possibly as well an enlarged or slightly changed version, in which the author took the opportunity of enlarging the novel by, say, adding a chapter or slightly revising some of its language. All of these items have, certainly, a common origin and relationship to a parent item. But they are not exact copies of the original text and thus are not exactly the same but at least are essentially the same as the parent work. Nevertheless, because they are essentially the same as the parent work, they too belong in the set of all items that are *manifestations* of the same *work*.

We can easily see how the definition works, since it provides a principle to organize each set that represents a single IE (each *work set*) in an integral way. However, the breadth of her notion, while providing inclusiveness, also encompasses ambiguity. Svenonius is not operational enough to build a cataloguing model. As the author says, "While this definition is formal, it is not wholly operational because it is not sufficiently constructive to identify

unequivocally all documents belonging to a work set" (2000, p.37). Another limitation of this notion of work is that it applies only to book materials, as Svenonius points out (2000, p.37). Later on, the FRBR document attempts to bridge this gap between *work* as an abstract concept and the call for operability, as well as to broaden the scope of *work* to non-book materials.

1.2. *Superwork*.

Svenonius' definition of *work* allows her to introduce a second concept: *superwork.* In the above transcribed quotation we saw two levels of relationships among items and a parent work related to how close they are to the original work's content—i.e., those items that are *exactly the same* and those that are *essentially the same*. Svenonius goes on to suggest, however, that there may be a third level of relationship, the relationship of those entities that have only a *likeness* to the original work. In this respect, Svenonius states that "a superwork may contain any number of works as subsets, the members of which while not sharing essentially the same information content are nevertheless similar by virtue of emanating from the same ur-work" (Svenonius, 2000, p. 38). If we bring together this new relationship, we can distinguish three levels of relationships from a work to other items, the first two of which constitute components of *a work set*: (i) a set of entities that are equivalent to the parent work (the closest), (ii) a set of items that are related to the parent work by

conveying essentially the same information (though revised, updated, abridged, enlarged or translated) and, (iii) a set of items that are related to but not included formally in the basic work set of the parent work. This last group includes other works that are far enough afield from the original—for example, adaptations, performances in another medium, commentaries—that while related to the original work are not really manifestations of it as *a work.* However, when the relationships to the original work-set are made clear, one has eventually created what Svenonius calls a *superwork.*

We can illustrate this additional group of related items, the *superwork,* for the above-mentioned case of Carsons McCullers. The new set will include variations of the content that make it *similar* to the actual work. In such set we could find a work by another author that constitutes a literary criticism of McCuller's work, for example, one written by a specialist that discusses the psychological (or sociological, or political) elements of the novel. An example of this kind of work is the master's thesis by Carlo Kilp *The heart is a lonely hunter : a study* (University of Texas at Austin, 1958). This master's thesis does not convey the same information that the original work does, but it is obviously related to the original. In a similar way, the movie directed by Robert Ellis Miller (based on the novel), belongs to this *superwork* set. Even though those additional items themselves might seem disparate, they are all related to the same basic original work, which is the artistic creation made by writer Carson

McCullers. Svenonius expresses this relationship between items within the set and the work as "the set of all documents descended from a common origin (superwork)" (Svenonius, 2000, p. 35).

1.3. *Bibliographic Family* of IEs.

Svenonius, in her definition of work focuses on all IEs that are related through a common origin. Richard Smiraglia develops the idea of groups of related items similar to Svenonius' ideas, but he does so in the form of a *family* of documents, introducing the notion of *parent work,* and giving to the general discussion a "genealogical" approach that makes it easier to "depict" or "map" the relationship among IEs.

In his treatise "The nature of 'a work'," Smiraglia (2001) describes *a work* as "the set of ideas created probably by an author or perhaps a composer, or other artist, set into a document using text, with the intention of being communicated to a receiver (probably a reader or a listener). A work may have many texts, and may appear in many documents" (Smiraglia, 2001, p.4). This definition contains new elements that add complexity and richness to the discussion. First, Smiraglia extends the limits of the nature of the items under study beyond text to the music realm. We are not only facing text-based messages, but also meanings and ideas made up of sounds. The nature of sounds introduces controversial issues to the nature of the content: does the content

consist of the sounds? Or are the sounds in fact the physical elements that convey the composer's content, which actually are higher feelings, ideas, etc? Furthermore, how does a performer's feeling modify the content or work made by the composer?

The second element that Smiraglia adds to the definition of work is the idea that the content is *communicated* from an "author or perhaps a composer" to a "receiver" (the "reader" or "listener"). This element of communication brings the user into the discussion for the first time. If we follow this thread, the set of items is not only related to the work by the common content they share, but also by the role that the user plays in the relationship. Does the work have the same "reader" or "listener" that each item within the set has? Is there any correlation between the "receiver" of an item and another "receiver" of a second item within the same set?

Although Smiraglia's issues need to be further explored, the value that he places in the notion of "receiver" is also developed in the FRBR report. In FRBR, the "user" (sometimes called the "receiver") becomes part of the justification for the creation and implementation of the proposed model. In its FRBR document, the IFLA Study Group claims to take a different approach to organizing information by analyzing each IE in the light of what the *user* (in Smiraglia's words: the receiver of the work's message) needs.

Smiraglia develops his own approach to conceptualizing "sets" of IE joined by a common work in the article "Bibliographic Families in the Library Catalog: a Qualitative Analysis and Grounded Theory" (co-authored with Gregory H. Leazer). In the section "Bibliographic Relationship and Families," Smiraglia and Leazer discuss their "genealogical" approach. "Any work…can serve as a progenitor for additional works" (Leazer and Smiraglia, 1999, p.192). Therefore, the work behaves as a "parent" that "produces" or "gives birth" to a set of IE. This point of view enforces even more the relationship between the work and its related items in the set of IEs.

All these ideas and findings are further developed by the IFLA Study Group that developed the FRBR model. Particularly, the concept of work will constitute the cornerstone for building its cataloguing model. On the other hand, the FRBR model will not take the conceptualization of *user* from Smiraglia. Instead, FRBR will bring the definition of user from the realm of database modeling techniques, invoking the idea of *user-centered approach* upon which these techniques are based. I will discuss these ideas in the next chapter.

Chapter Two: The Functional Requirements for Bibliographic Records (FRBR) Model

2.1. Goal of the FRBR model.

The goal of the document *Functional Requirements for Bibliographic Records* (FRBR), written by the IFLA Study Group on the Functional Requirements for Bibliographic Records, is to place various notions or concepts related to work and item (copies) in an integrated system for creating surrogate records in cataloguing. FRBR not only adopts the idea of work from Lubetzky but also borrows a framework theory from the database management systems framework, *the entity-relationship data model* used by designers to develop database management systems (DBMS)

The purpose of bringing DBMS as a baseline component of the FRBR model is to address the analysis and creation of surrogate records from the standpoint of users' needs. As the document says:

> The study uses an entity analysis technique that begins by isolating the entities that are the key objects of interest to users of bibliographic records. The study then identifies the characteristics or attributes associated with each entity and the relationships between entities that are most important for users in formulating bibliographic searches, interpreting responses to those searches, and "navigating" the universe of entities described in bibliographic records. (IFLA, 1998, p.4)

From this quotation we are told that FRBR will analyze the entities as far as the user is concerned. However, the report provides no definition of *user* or *users*.

Instead, it provides only a brief definition that says that the term *user* encompasses a "broad spectrum" of individual and institutions, "including not only library clients and staff, but also publishers, distributors, retailers…" (IFLA, 1998, p.4). But other than naming those users, the document does not define their profile, motivations, information seeking behavior, etc. I think that this lack of exploration of the concept of user and its general implications in cataloguing as well as its particular operation in the FRBR model is a shortcoming of the model.

In reality, database modeling, the science that became the framework for the document, has a record for analyzing both users' needs and information seeking behavior through a wide range of testing tools such as *usability tests.* Employing these tests as instruments, database developers study users' needs with an objective methodology, disregarding any assumptions that the developer may have about who those users are and what they need. Multidisciplinary teams conduct most of these techniques. As Mayhew (1999) suggests, the range of techniques to determine users' profiles may vary, but all include basic steps such as identifying users' categories, determining relevant users' characteristics, and gathering data (Mayhew, 1999).

In a similar fashion, cataloguers will have to develop or borrow tests and procedures to analyze users' behaviors and needs while seeking information. The IFLA Study Group would have taken a great step forward if they had

conducted a test instrument to study library catalogue's users. The above quotation states that the model functionalizes only those attributes and relationships of an IE that are important to users. But, are they depicting an "actual" user or an ideal user from the cataloguer's point of view? Where are the usability tests and the users' profile studies? The declarations about users' needs are not founded in any study about the nature of those users. Undoubtedly, this area of study presents challenges for further research.

In the following sections, I will first introduce the framework theory of the report, the *entity-relationship data model* (ERDM), and then I will explain the FRBR model.

2.2. Database System Modeling

"Complex design activities require conceptual simplicity to yield successful results." (Rob, P. & Coronel, C., 2002, p.36).

2.2.1. The Entity-Relationship Data Model

The origin of the Entity Relationship Data Model (ERDM) goes back to the 1970s, when database developers were seeking a way to depict database entities and their relationships in a graphical fashion. The context for the development of ERDM was the discovery and wide implementation of relational databases. One necessity for the latter was to find a graphical way to depict relationships among tables in such databases (i.e. relationships among database objects). The

basic standpoint of research was that it is easier to examine structures in a graphic array rather than in a textual, linear way. In this context of research, Peter Chen developed and introduced a tool called the *entity relationship database model,* ERDM. The model serves as a bridge between real-world objects and the database representations of those objects (Rob, P. & Coronel, C., 2002, p.109). ERDB is also a powerful communication tool among database designers, programmers and users. Since the main goal of any database system is to satisfy users' needs, communication among programmers and users is essential to keep the design focused.

According to the ANSI/SPARC[1] there are three kinds of data modeling, which correspond to three levels of abstraction from reality to database design: a conceptual model, an internal model and an external model. The first of these, the *conceptual model*, consists of an overall graphical picture of the design. It is a "bird's eye view" that roughly shows main components and their relationships. It constitutes an abstract way of representation. At a more concrete level, programmers develop an *internal model*, which consists of a prototype for implementation. It includes general definitions of components, and relationships in the model are more closely drawn focusing on a particular database management system. The final level is the *external model,* which allows programmers to concentrate on particular groups of users. In this stage, the

developer takes some parts of the internal model and works along with the end-users of the system to identify users' needs. This level is the most detailed depiction of the database system (Rob, P. & Coronel, C., 2002, p.23).

2.2.1.1. Elements within the model.

The structure of the ERDM models is based on three basic elements: *entity, attribute*, and *relationship*.

a. *Entity: entity* is "a person, place, or thing about which data are to be collected and stored." (Rob, P. & Coronel, C., 2002, p.36). In a database system modeled for managing students' registration to classes in an education institute[2], for example, entities will be *student*, *professor*, *course* and *class*.
b. *Attributes:* each entity has *attributes*, which are particular characteristics belonging to it, such as *student's name* and *student's ID number*, *class's size*, or *professor's phone number*.
c. *Relationship:* finally, a *relationship* is an association between entities (Rob, P. & Coronel, C., 2002, p.124). In terms of database modeling, relationships are expressed in a simple sentence like "a professor *teaches* a class" or "a class *contains* students."

[1] American National Standards Institute/Standards Planning and Requirements Committee

2.2.1.2. Relationships among elements.

Relationships are fundamental elements in ERDM. The consistency of the system will strongly depend on the level of accurateness that relationships have among entities. Also, the concept of relationship as developed by ERDM will be closely followed by the FRBR document. In this section, I will discuss the relationships that are taken by FRBR model: *participation* and *type*.

Relationship participation. ERDM determines a relationship *participation* as either *optional* or *mandatory*. An *optional* relationship is one in which the generation of an occurrence for an entity does not determine the generation of a related occurrence in its linked entity. In the example mentioned above, the existence of an occurrence in the entity *professor* does not determine that the linked entity *class* will have one (since a professor may or may not have a class during a particular semester). This constitutes an optional participation relationship. Conversely, an occurrence under the entity *class* will require the corresponding occurrence in the entity *professor* since every class has to have a professor. In this situation, therefore, a *mandatory* relationship exists, since "one

[2] I took the educational institute example from Rob, P. & Coronel, C., 2002, third chapter, within their discussion of the model.

entity occurrence *requires* a corresponding entity occurrence in a particular relationship" (Rob, P. & Coronel, C., 2002, p.130).

Relationship type: connectivity and cardinality. There may be three kinds of relationships according to the number of occurrences that the relationship determines between the entities. They are *one-to-one, one-to-many, and many-to-many*. A *one-to one* relationship is exemplified for entities *professor* and *IDNumber* since each professor has only one IDNumber and vice versa. This means that a single occurrence in the first entity corresponds to another single occurrence in its related entity.

One-to-many relationships appear more often. They occur when the occurrence generated under one entity is related to many occurrences in its linked entity. A professor can teach many classes, for example. The relationship is represented by the symbols (1, ∞), where ∞ stands for "many occurrences."

A *many-to-many* relationship is, in fact, a wrong relationship within the conceptual model. A relationship of this kind (such as the situation of a student who enrolls in several classes and, conversely, a class that has many students) may be detected but, for the sake of consistency in the system, must be cleaned out by building a *bridge* among entities. A bridge is a third entity between the entities. Bridging is an action and in the example here the bridge will be *enroll*, and will join the entities *students* and *class* by standing in between them.

What has been described so far is the kind of relationship (one-to-one, etc.) involved in connectivity. *Cardinality* further specifies *connectivity* by describing the number of occurrences involved. For example, cardinality (1,1) corresponds to a one-to-one kind of relationship. If the school in our example, however, has 20 classes, then the connectivity relationship changes to *one-to-many* among the entities *professor* and *class* and thus will be expressed as *cardinality* (1,20).

The FRBR model will take advantage of all these relationships to give consistency and comprehensiveness to its model. In the section ahead that describes relationships among such entities as *work* and *manifestation*, we will discuss how operational definitions such as these provide the FRBR document with a great source of tools to implement a coherent system.

In the following section, we will discuss how FRBR takes all those elements of ERDM and integrates them to the notions described in chapter one, i.e. *work* and *superwork.* Basically, what the FRBR model does is to turn Svenonius' operational definition of *work* into what we defined as an *entity* under the ERDM framework. However, *work* won't be a raw entity among others; rather, it will be the main reference point for other kinds of entities. The model then will build a hierarchical relationship among entities in relation to the entity *work.*

2.3. The FRBR model[3]

The FRBR document outlines the three above-mentioned key elements for any database model system: *entities, attributes* of those entities, and *relationships*. Also, the model relates attributes and relationships to *users' tasks*. Finally, in its last section, the document determines the basic cataloging requirements for National Bibliographic Records regarding the mapped users' tasks.

In this report, I will focus on section one through six of FRBR. These sections are the ones that model the system and incorporate the notion of a work as discussed in chapter one. The modeling of users' tasks, even though important to the whole system, exceeds the scope of this paper. Also, the "Basic Requirements for National Bibliographic Records," the last chapter of the FRBR report, go beyond the discussion of *superwork* since such chapter attempts to solve the problem of a minimal level of fields in a record for worldwide bibliographic exchange.

2.3.1. Entities

The FRBR model takes the operational concept of entities from ERDM that we discussed and classifies its own entities within three groups:

[3] I took all elements for this discussion from the *Final Report* document by the IFLA Study Group on the Functional Requirements for Bibliographic Records (see References page)

- *GROUP 1 ENTITIES. The artistic or intellectual product itself*: what we traditionally meant by "the book," or "the music record." The IEs comprised in this group are those related to the essential unit of information sought by users: novels, songs, maps, websites, etc. The group includes four database system entities: *work, expression, manifestation* and *item.* The sequence of entities follow the level of abstraction that they have, i.e. from the most abstract -- which is the *work*-- to the least abstract, physically defined level, which is the *item* entity.
- *GROUP 2 ENTITIES. The individual/s of institution/s that created the artistic or intellectual product.* Cataloging rules and procedures associate this group of entities with the notion of "authorship" in the case of individuals and "emanation" in the case of corporate bodies. However, FRBR goes beyond this concept while including those responsible for "the physical production and dissemination, or the custodianship of the entities in the first group" (FRBR, 1998, p. 13). The entities in this group are: *person* and *corporate body.*
- *GROUP 3 ENTITIES. The categories in which the work can be classified*: they include *subject, object, event* and *place.* This categorization of IE provides users with an intuitive way of

browsing the library catalogue. *Subjects* such as "literature" or "biology" are part of the first kind of categories. *Objects* can be animated or unanimated, current or historical, and so on. Examples of objects are sculptures, stones, art craft, monuments, etc. Finally, both *event* and *place* are historical or geographical qualities of IE with which the *work* is worth being classified. All four of these categories (*subject, object, event* and *place*) become entities only when one or more works discuss them. For example, the twin towers in New York became an entity related to the terrorist attack of September 11 when one or more IEs were generated to commemorate them. This universe is composed of pictures, physical reproductions at scale, tales about the tragedy of the plane crash, etc.

For the scope of this essay, I will discuss only the first group of entities, i.e. those entities that encompass the artistic or intellectual product itself. I limit my analysis to this group because it focuses on the notion of *work* and its related entities.

2.3.1.1. Group 1 Entities

Work. Work is the most important and also the most elusive concept in the FRBR model. We discussed the idea of work in chapter one from the point of

view of developments in cataloguing. In this section, I will explain the FRBR's approach.

FRBR briefly defines work as "a distinct intellectual or artistic creation" (FRBR, 1998, p. 16). Such a creation should be conceptualized apart from any physical realization, appearance or container of it. FRBR draws this differentiation between abstract and concrete in a very subtle, though strongly operative, way.

To better understand the entity *work* in the FRBR model, we may say that, for example, in the tale "El Aleph," by author Jorge Luis Borges, the work consists of the ideas, characters and events that the narration conveys and which are essentially abstractions made up by Borges' creative talent. Also, these creations are carried by sequences of letters, words, sentences, and so on that the artist put together. But such sequences are not to be called "the work" because they remain at the physical level (i.e. the physical written printing in a paper). If those sequences of words were actually the work itself, we would have to face many problems in defining the set of related entities that make up the work called *El Aleph*. One of these problems is how to relate a translation of the tale into, say, English. Since the chain of letters, words and sentences changes, we would have to conclude that a translation is a separate, different work. But separating related entities in this way is not only inappropriate because it breaks up a relationship that does exist, but also goes against the goal of the FRBR

model, i.e. to allow catalogs to develop the idea of separated works as sets of items related by a common origin and with essentially the same content. Furthermore, if we go up in this consideration of *work* as involving the physical level of words, a translation of that tale into a non-alphabetic language such as Arabic or Russian would be tantamount to a new set of items different from the work *El Aleph* (original in Spanish), their versions in alphabetic languages (such as English), and other versions in no-alphabetic languages, i.e. translations into Arabic, Hindi, Chinese, etc. We will end up having three different "works" with a set of items depending on each one.

Expression. The so-described *work* is realized through an *expression.* "An expression is the specific intellectual or artistic form that a *work* takes each time it is 'realized'" (FRBR, 1998, p. 18). This realization, however, does not imply a physical item (we are still far from the physical level). Rather, expression involves changes in the way in which the work is conveyed due to changes in: (i) *form*, such as from alpha-numeric notation to spoken word, (ii) "intellectual conventions or instruments that are employed to express the work" (FRBR, 1998, p. 19).

The notion of *expression* helps us to understand the above-mentioned problem of translations. In the case of Borges, the sequence of sounds constitutes the *expression* of his ontological ideas about the universe and its

"flow" through history. *Expression,* in consequence, goes a step further from abstraction to concreteness. But, it is not yet fixed in total concreteness, which is reserved for items alone.

As already mentioned, a clear case that depicts the differences between expressions and works is the case of translations. When a work is translated from one language to another, the final outcome is two *expressions* of the same work: the one in the original language and the one translated. The following example, quoted from FRBR, shows the difference:

W_1 Ellwanger's *Tennis—bis zum Turnierspieler*
E_1 the original German text
E_2 the English translation by Wendy Gill

W_2 Franz Schubert's *Trout quintet*
E_1 the composer's score
E_2 a performance by the Amadeus Quartet and Hephzibah Menuhin on piano
E_3 a performance by the Cleveland Quartet and Yo-Yo Ma on the cello

(Source: FRBR, 1998, p. 19: **W**=work; **E**=expression).

The idea of *expression* allows us to join together a set of items, each of which is related to the same parent-IE: the work. The instrumental power of this concept consists of gathering bibliographic entities that otherwise woul be spread out. In fact, this spread-ability is the norm of our current catalogs' model. For example, searching for all related items that are in reality *expressions* of the *Iliad*, involves a quite long sequence of steps for the user. He has to search by the title

and then carefully examine the different items retrieved in order to differentiate among first, second, etc. editions, publications issued in different places (such as countries), editions with comments, movies, music created from the text, and so on. Also, the user might have to figure out a way of finding the work translated into other languages, since the title will be in that language, at least in cases where uniform titles have not been used.[4] This can be done through an author search and then selecting the particular work and its translations, but usually that requires some knowledge of the language into which the work was translated ("La Ilíada" in Spanish, for example). All these steps are quite complicated. Having a depiction of *expressions* (like editions and translations) related to the *work* (*The Iliad*) would constitute an important step forward to better satisfy users' needs.

Even with these positive aspects, we will discuss in the conclusions to this report that the notion of expression is controversial. Edward O'Neill (2002), for example, points out that expression in fact constitutes an unnecessary entity that can be taken out for the sake of simplicity, replacing it by adding more attributes to the manifestations (O'Neill, 2002, 158). We will discuss this issue later on.

[4] In traditional cataloguing, uniform titles, while widely used in the case of very classic works such as *The Bible*, are not employed in all cases.

Manifestation. As the FRBR document discusses, "when a *work* is realized, the resulting *expression* of the *work* may be physically embodied on or in a medium such as paper, audio tape, video tape, canvas, plaster, etc. This physical embodiment constitutes a *manifestation* of the work" (FRBR, 1998, p. 20). Therefore, we face here a much greater physical level in the entity's sequence. The *expression* of the *work* titled "*Iliad*" has several *manifestations* according to all physical sets of bibliographic items produced. The category *Manifestations* includes things such as page layout for printed works, tape records for music items, dimensions for objects, etc.

The following example helps to understand the different levels of entity:

W_1 Jean Jolivet's *Vrai description des Gaules…*
E_1 the cartographer's original rendering
M_1 the map issued in 1570
M_2 a facsimile reproduction published in 1974

(Source: FRBR, 1998, p. 22: **W**=work; **E**=expression; **M**=manifestation).

However, this level of entity still is not the concrete, singular item that we hold in our library. As FRBR says, "as an entity, manifestation represents all the physical objects that bear the same characteristics, in respect to both intellectual content and physical form" (FRBR, 1998, p. 20). Manifestation, therefore, is an abstract instance that bridges the abstract level to the concrete one but still belonging to the former. For example, the set of manifestations of a work published in CD-Rom join together all items that realize the expression of the

work through that physical embodiment. The concept of manifestation joins such items, but that concept essentially does not encompass a physical state of existence.

Item. The last entity corresponding to the first group is *item*. *Item* is the actual physical (i.e. totally concrete) informational object that we hold in our library, stored in our computer, accessed through the Internet by a URL address, and so on. An *item* is a single physical object that conveys the *manifestation* or a part of it. For example, we may hold "Iliad" in one single book, or in a two-volume edition. These two volumes are two *items* that belong to one single *manifestation*. Again, an example taken from the FRBR text:

W_1 Ronald Hayman's *Playback*
E_1 the author's edited text for publication
M_1 the book published in 1973 by Davis-Poynter
I_1 item autographed by the author

(Source: FRBR, 1998, p. 23: **W**=work; **E**=expression; **M**=manifestation; **I**=item).

2.3.2. Attributes

Going back to the ERDM, we recall that each entity has *attributes.* These attributes are characteristics of interest for the database manager and constitute also the actual data that users deal with in the system. In the example handled in

section 2.2.1 (a database for managing students' registration to classes in an education institute), we may have a set of attributes for the entity PROFESSOR such as "FirstName," "LastName," "PhoneNumber" and so on. While performing a search in a database, we never look it up for the entity "professor" as such. Rather, we identify each professor by his or her characteristics or attributes, such as his or her last name.

In a similar way, each entity in the FRBR model has a set of related attributes. For example, the entity *work* has twelve attributes, among them *title, form, date,* and so on. Similarly, the entity *expression* has a set of twenty-five attributes such as *title of the expression, form, extensibility, context,* etc. In turn, *manifestation* as an entity encompasses thirty-eight attributes; some of them are *title, statement of responsibility, edition, publisher,* and *form*. Finally, the entity *item* carries nine attributes such as *item identifier, provenance* and *condition.* Those attributes are fundamental elements in searching and identifying a work or any of its expressions and manifestations. As in other ontological levels, we recognize entities through attributes since an entity's essence is elusive to grasp in a natural and straight way.

Each entity has many attributes. For the purpose of this report, the exhaustive list of such attributes will be given in chapter three when developing the set of surrogate records for the work "Rayuela," by Julio Cortázar. As mentioned in the Introduction, this chapter describes the set of holdings at the

Benson Latin American Collection, University of Texas at Austin. The reason for discussing entities' attributes in that chapter is that such list is strongly related to the nature of each entity. Thus, it will make more sense to describe such lists in concrete examples, such as the items that make up the work *Rayuela.*

2.3.3. Relationships

The FRBR document spends an important section explaining the *relationship* among elements in the model. *Relationships* are defined as "the vehicle for depicting the link between one entity and another, and thus the means of assisting the user to 'navigate' the universe that is represented in a bibliography, catalogue, or bibliographic database" (Source: FRBR, 1998, p. 56). The importance of *relationships* in the model should not be understated. The power of the accuracy of the IE retrieved on the search (i.e. the *precision* of our database results, in terms of pattern matching algorithms[5]) depends on how accurately and relevantly these relationships between entities are established. As FRBR points out, well-defined relationships help the user to smoothly navigate among IEs. Thus, the way in which relationships are depicted in the surrogate records' system model will help or hinder the quality of navigation in the catalogue.

[5] For further discussion about these concepts see Rosenfeld and Morville (2002)

The FRBR report also points out that relationships among entities have to be clearly stated in the catalogue because, in turn, they may not be clearly stated in the bibliographic item itself. For example, an edition that the principal source of information for description (i.e. the title page) names as *augmented* may be in reality *revised* or *modified.* This confusion constitutes a real problem in the set of items that I have chosen for a demonstration—the items that make up the work, *Rayuela.* In Spanish, there is no technical difference between the concepts of *edition* and *printing*—both are named edition ("edición"). Therefore, when an item called *Rayuela* has the legend "segunda edición" (second edition) in the title page, it means that the item is actually the second *printing* of the work[6]. Through this concrete example we can see how the terminology used in different publishing cultures differs and is reliant on common usage. For this reason, FRBR discusses and defines each single relationship among entities. I think this level of detail is one of the strong points of the FRBR document.

FRBR depicts relationships only at the higher level of abstraction in the model—group 1 entities. There are two kinds of relationships: (i) hierarchical relationships depicted between *work, expression, manifestation* and *item*, and that make up one work, (ii) other relationships, both hierarchical and horizontal,

[6] In Latin American publishing, the 2nd, 3rd, etc, editions will be highlighted indicating the kind of modification that the edition involves, such as augmented or revised, along with the edition

among entities that represent different works. Following the scope of this report, I will focus on relationships (i), hierarchical relationships among entities that make up the work, since they develop the links that allow the operation of the concept of *work.* How the vertical line of associations is portrayed will determine whether or not the system "maps" the work and its elements in a system focused on delineating all the expressions and manifestations of a work. I will explore the *relationships* in the FRBR model from this approach.

2.3.3.1. Hierarchical relationships

The relationships among *work, expression, manifestation* and *item* are as follows:

- A work is realized through an expression,
  - Or an expression is a realization of a work.
- An expression is embodied in a manifestation,
  - Or manifestation embodies an expression.
- A manifestation is exemplified by an item,
  - Or an item exemplifies a manifestation.

(Source: FRBR, 1998, Section 3.1.1 and Figure 3.1., p. 12-13)

The FRBR document clearly states: "each of [these] three primary relationships [realization, embodiment, and exemplification] is unique and operates between only one pair of entities in the model" (FRBR, 1998, p. 58). When moving

---

number.

upward in the model, the uniqueness and verticalness of such relationships ensures consistency, i.e. a expression will correspond to one and only one work, a manifestation will correspond to one an only one expression and, finally, an item will correspond to one and only one manifestation.

However, if we go in the reverse direction through the relationships, the correspondence is not unique anymore. This means, for example, that a work can be realized in several expressions, while an expression is such only in relation to a single work. In terms of ERDM modeling, we say that entities *work* and *expression* have a *connectivity* of *one-to-many* and that their *cardinality* is $(1, \infty)$, where ∞ stands for "many occurrences" for the entity[7]. This relationship type gives consistency to the system, being highly recommended by ERDM modeling specialists (Rob, P. & Coronel, C., 2002, p.126).

2.3.3.1. Other relationships

Aside from the foregoing relationships, other relationships (both vertical and horizontal) also exist within the model. For example, the *work-to-work* relationship associates horizontally the higher level of entities, the work, in group 1 entities. Following the *work-to-work* relationship scheme, one work can be related to another work by seven categories: *successor, supplement, complement, summarization, adaptation, transformation, and imitation* (FRBR,

1998, p. 65). The first category, *successor,* relates a work that is subsequent of another. For example, Julio Cortázar, the author of the work that I will model in chapter three, wrote the novel *62/ Modelo para Armar* (translated to English as "*62: A Model Kit*"). This novel in fact constitutes chapter number sixty-two of the work *Rayuela.* That is the reason for its title. Just as a literary game, he did not write and insert chapter sixty two in between the other chapters in the novel. He later published this chapter as an independent work. As we can deduce, establishing this kind of relationship allows a user to link the work *62/ Modelo para Armar* and the work *Rayuela.*

The kind of relationship *whole/part* constitutes an example of a vertical relationship. This relationship, which is given within all entities in group 1, establishes links from a work, expression, manifestation, or item, to either dependent of independent entities of the same nature (i.e., whole/part work-to-work relationships, whole/part expression-to-expression relationships, etc.). An example of whole/part work-to-work relationship may be the section of a work (such as a chapter) with the work itself (dependent part) or a journal article with the journal (independent part). In this way, entities are related in the same level without losing their original independence.

At this point I should explain that, when these relationships are spelled out, one is in reality entering the realm of Svenonius' *superwork.* However, the

[7] See section 2.2.1.2. *Relationships among elements*

FRBR model does not represent or discuss that concept. Therefore, I have not used it when applying FRBR to the novel used here as a test.

In the following chapter I will apply the FRBR model for the work *Rayuela*, by Julio Cortázar, by developing surrogate records comprising the entities and attributes described above. As far as relationships are concerned, I will deploy only hierarchical relationships, and within group-1 entities. This focus will provide with a more consistent analysis to the notion of *work*.

Chapter Three: an Application of the Model: *Rayuela*, by Julio Cortázar: Surrogate Records for the Holdings of this Novel at BLAC-UT Austin

In this section I will deploy the model proposed by FRBR in a particular set of items held at the Benson Latin American Collection (BLAC), University of Texas at Austin. This set consists of all items of the novel "Rayuela" by the Argentinean author Julio Cortázar. The modeling will allow depicting all editions of the novel in a comprehensive way, showing the relationships among all IEs of the novel held at BLAC. This "mapping" will allow the user to quickly identify the item needed and confront it to other items within the set.

1. Attributes

As mentioned before, the description of attributes is fundamental for the identification and retrieval of IEs. The FRBR model lists all possible attributes or characteristics that a user can need for identifying an item and relating it to a set belonging to a common *work*. Each one of the attributes themselves has been extensively used in library cataloguing; therefore, the FRBR document does not discuss them in detail. The following section lists all attributes for group-1 entities in FRBR, i.e. *work, expression, manifestation* and *item*.

It is important to mention here that we will not see in the following lists such elements as *author* or *corporate bodies.* In the FRBR model, both *persons*

(as authors) and *corporate bodies* (as organisms from which a work emanates) belong to group 2 entities.[8] In the frame of FRBR, group 1 entities (*work, expression, manifestation* and *item*) and group 2 entities (*person* or *persons* and *corporate bodie/s*) connect through relationships of *responsibility*, for example: a work *is created by* a person or person (or corporate bodie/s), an expression *is realized by* a person or persons (or corporate bodie/s), and so on. Although the depiction of these relationships is very important, FRBR does not discuss them as far as it does with the relationships within group 1 entities. Also, I have a single author for the set of all items I will analyze, which makes the relationship straight. For these reasons, and to keep the scope of my report, I will not deploy the relationship among group 1 entities (*work, expression, manifestation* and *item*) and group 2 entities (*person* or *persons* and *corporate bodie/s*) for the novel I have chosen.

1.2. Attributes for group-1 entities

1.2.1. Attributes of a *Work.*

Title of the work
Form of the work
Date of the work
Other distinguishing characteristics (those picked up to differentiate the work from another work that has the same title)
Intended termination
Intended audience
Context of the work

[8] See section 2.3.1. Entities

For musical works: Medium of performance
Numeric designation
Key
For cartographic works: Coordinates
Equinox

1.2.2. Attributes of an Expression

Title of the expression
Form of expression
Date of expression
Language of expression
Other distinguishing characteristics (those used to differentiate the expression from another expression related to the same *work*)
Extensibility of expression (applied when the expression is issued in several parts)
Revisability of expression (used for drafts in monographic expressions and expressions that are updated)
Extent of the expression
Summarization of content
Context for the expression
Critical response to the expression
Use restrictions on the expression
For serial expressions: Sequencing pattern
Expected regularity of issue
Expected frequency of issue
For musical notation expressions: Type of score
For musical notation and sound recording expressions: Medium of performance
For cartographic image/object expressions: Scale
Projection
Presentation technique
Representation of relief
Geodetic, grid and vertical measurement
For remote sensing image expressions: Recording technique
Special characteristics
For graphic or projected image expressions: Technique

1.2.2. Attributes of a Manifestation

Title of the manifestation
Statement of responsibility
Edition/issue designation
Place of publication/distribution
Publisher/distributor
Date of publication/distribution
Fabricator/manufacturer
Series statement
Form of carrier
Extent of the carrier
Physical medium
Capture mode
Dimensions of the carrier
Manifestation identifier
Source of acquisition/access authorization
Terms of availability
Access restrictions on the manifestation
For printed books manifestations: Typeface
Type size
For hand-printed manifestations: Foliation
Collation
For serial manifestations: Publication status
Numbering
For sound recording manifestations: Playing speed
Groove width
Kind of cutting
Tape configuration
Kind of sound
Special reproduction characteristics
For image manifestations: Color
For microform manifestations: Reduction ratio
For microform or visual projection manifestations: Polarity
Generation
For visual projection manifestations: Presentation format
For electronic resource manifestations: System requirements
File characteristics
For remote access electronic resource manifestations: Mode of access
Access address

1.2.2. Attributes of an Item

Item identifier (call number, Library's inventory number)
Fingerprint
Provenance of the item
Marks/inscriptions
Exhibition history
Condition of the item
Treatment history (in case of restored items)
Scheduled treatment
Access restrictions on the item (such as in-library use item)

This long list, while comprehensive, can be confusing. FRBR does not go deeply into the description of the attributes and the differences among them. For example, the list of attributes for the three first entities starts with the *title* of such entity (i.e., title of the work, title of the expression and title of the manifestation). However, the definition for each one does not differ among them. FRBR defines the attribute *title of the work* as "the word, phrase, or group of characters naming the *work*" (FRBR, 1998, p. 33). When we look for the definition of *title of the expression,* we find that such title is "the word, phrase, or group of characters naming the *expression*" (FRBR, 1998, p. 36). Finally, the definition of *title of a manifestation* says that "the title of the *manifestation* is the word, phrase, or group of characters naming the *manifestation*" (FRBR, 1998, p. 41). There are no further definitions or discussion about the differences among these types of titles.

In the particular case of *Rayuela*, for example, the task of detailing attributes was confusing and, at the same time, cumbersome. For example, in Spanish the title remains the same through all the three occurrences for the *title* attribute, i.e. title of the work, title of the expression, and title of the manifestation. In the case of translations, we have, say, the work *Rayuela*, then the expression in English titled *Hopscotch*, and finally the manifestation's title, which happens to be the same as the expression's one. This repetition seems to be awkward; it also adds arguments for criticizing the model. O'Neill (2002) for example, reasonably argues that the entity *expression* can be removed to avoid unnecessary duplication of levels in the FRBR model. We will discuss these issues later on in the conclusion of the report.

2. The FRBR Model Applied to *Rayuela*, by Julio Cortázar

To analyze the work *Rayuela*, I will deploy only the attributes related to each particular entity within the group 1 (*work, expression, manifestation* and *item*). A very important observation is that, to make the set more readable, I will not transcribe attributes that repeat from one surrogate record to another in the same set of entities. For example, there are several editions of the work published by "Editorial Sudamericana" (Buenos Aires). I will not transcribe the attribute *place of publication* in an *item* when it is the same than it was for the *manifestation* that precedes it.

A very interesting issue that the holdings for this work at the Benson Latin American Collection raise has to do with what is the actual work. For this novel the Benson has both the original manuscript of the work and also the first issued edition of the work. Is the actual work best expressed in the form of the manuscript or in the form of the first issued edition? Here, I have concluded that the manuscript is a draft (with all strikethrough, sub and superscripts and notes that a draft has), and that the work set for the novel started to develop only after the first printed item was issued. Thus, I will take this first printing as the basic expression of the *work.* It is interesting, nonetheless, to note that the problem could be further investigated because no definitive researches of these relationships have been made.

It is also very important to explain that, technically, the use of "edición" (*edition* in English) is not the same in Spanish than its related term in English. "Edición" in Spanish merely means *printing.* In English, it means a new impression of the work without modifications. Therefore, when we have manifestations that say "primera edición" (*first edition*), "segunda edición" (*second edition*) and so on, this means first, second "printing" and nothing more than that. However, since FRBR does not recognize this distinction, each differently designated edition here has been considered a different manifestation. More will be said about this in the conclusion to this report.

2.1. The set of IEs for *Rayuela*

WORK.
Attributes of the Work.
Title: Rayuela
Author: Julio Cortázar
Form: novel
Context of the work: existentialism (philosophical); dictatorship in Argentina (historical).

EXPRESSIONS: SET # 1: ORIGINAL WORK
Attributes of the expression
Title: Rayuela
Language: **Spanish Versions**
Form: Written word

Manifestations
Manifestation 1
Attributes of the Manifestation 1
Title: Rayuela
Statement of Responsibility: Julio Cortázar
Edition: 1st ed.
Place of Publication: Buenos Aires
Publisher: Sudamericana
Date of Publication: 1963

Items from Manifestation 1
Item # 1
Attributes
Item identifiers:
Call Number: PQ 7797 C7145 R3 1963
Provenance of the item: Presented by Beverly Gibbs

Manifestation 2
Attributes of the Manifestation 2
Title: Rayuela

Statement of Responsibility: Julio Cortázar
Edition: 9th ed.
Place of Publication: Buenos Aires
Publisher: Sudamericana
Date of Publication: 1968

Items from Manifestation 2
Item # 1
Attributes
Item identifiers:
Call Number: PQ 7797 C7145 R3 1968 LAC
Inventory Number:. 3002167624
Provenance of the item: Presented by Beverly Gibbs

Manifestation 3
Attributes of the Manifestation 3
Title: Rayuela
Statement of Responsibility: Julio Cortázar
Edition: 11th ed.
Place of Publication: Buenos Aires
Publisher: Sudamericana
Date of Publication: 1969:

Items from Manifestation 3
Item # 1
Attributes
Item identifiers:
Call Number:. PQ 7797 C7145 R3 1969 LAC
Inventory Number:. 2114543110
Provenance of the item: Presented by Barnes Lathrop

Manifestation 4
Attributes of the Manifestation 4
Title: Rayuela
Statement of Responsibility: Julio Cortázar
Edition: 17th ed.
Place of Publication: Buenos Aires
Publisher: Sudamericana
Date of Publication: 1974

Items from Manifestation 4
Item # 1
Attributes
Item identifiers:
Call Number: PQ 7797 C7145 R3 1974 LAC
Inventory Number: 2013274821

Manifestation 5
Attributes of the Manifestation 5
Title: Rayuela
Statement of Responsibility: Julio Cortázar
Edition: 18th ed.
Place of Publication: Buenos Aires
Publisher: Sudamericana
Date of Publication: 1975

Items from Manifestation 5
Item # 5.1
Attributes
Item identifiers:
Call Number: PQ 7797 C7145 R3 1975 LAC
Inventory Number: 2013274832

Item # 5.2
Attributes
Item identifiers:
Call Number: PQ 7797 C7145 R3 1975 LAC COPY 2
Inventory Number: 3020512746

Manifestation 6
Attributes of the Manifestation 6
Title: Rayuela
Statement of Responsibility: Julio Cortázar
Edition: 19th ed.
Place of Publication: Buenos Aires
Publisher: Sudamericana
Date of Publication: 1976

Items from Manifestation 6
Item # 1.
Attributes.
Item identifiers:
Call Number: PQ 7797 C7145 R3 1976 LAC
Inventory Number: 2013274843

Manifestation 7
Attributes of the Manifestation 7
Title: Rayuela
Statement of Responsibility: Julio Cortázar
Edition: 21th
Place of Publication: Buenos Aires
Publisher: Sudamericana
Date of Publication: 1977

Items from Manifestation 7
Item # 1
Attributes
Item identifiers:
Call Number: PQ 7797 C7145 R3 1977B LAC COPY 2
Inventory Number: 2102012591

Manifestation 8
Attributes of the Manifestation 8
Title: Rayuela
Statement of Responsibility: Julio Cortázar
Place of Publication: Santiago (Chile)
Publisher: Seix Barral
Date of Publication: 1985

Items from Manifestation 8
Item # 1
Attributes
Item identifiers:
Call Number: PQ 7797 C7145 R3 1976 LAC
Inventory Number: 2101588353

Manifestation 9
Attributes of the Manifestation 9
Title: Rayuela
Statement of Responsibility: Julio Cortázar
Place of Publication: Barcelona
Publisher: Edhasa/Sudamericana
Date of Publication: 1977

Items from Manifestation 9
Item # 1
Attributes
Item identifiers:
Call Number: 7797 C7145 R3 1977C LAC

Expressions: set # 2
Attributes of the expression.
Title: Hopscotch
Language: **English Translations**
Form: Written word

Manifestations
Manifestation 1
Attributes of the Manifestation
Title: Hopscotch
Statement of Responsibility: Julio Cortázar
Statement of Responsibility for the Translation: Gregory Rabassa
Place of Publication: New York
Publisher: Pantheon Books
Date of Publication: 1966

Items from Manifestation 1
Item # 1.1
Attributes
Item identifiers:
Call Number: PQ 7797 C7145 R3 1977B LAC, copy 1.
Inventory Number: 3004878532.

Item # 1.2
Attributes.
Item identifiers:
Call Number: PQ 7797 C7145 R3 1977B LAC, copy 2.
Inventory Number: No inventory number found.

Manifestation 2
Attributes of the Manifestation 2
Title: Hopscotch
Statement of Responsibility: Julio Cortázar
Statement of Responsibility for the Translation: Gregory Rabassa
Place of Publication: New York
Publisher: Pantheon Books
Date of Publication: 1986

Items from Manifestation 2
Item # 1
Attributes
Item identifiers:
Call Number: PQ 7797 C7145 R313 1987 LAC.
Inventory Number: 2105215948.

Manifestation 3
Attributes of the Manifestation 3
Title: Hopscotch
Statement of Responsibility: Julio Cortázar
Statement of Responsibility for the Translation: Gergory Rabassa
Place of Publication: New York
Publisher: Avon Books
Date of Publication: 1975

Items from Manifestation 3
Item # 1
Attributes
Item identifiers:
Call Number: PQ 7797 C7145 R313 1975 LAC.
Inventory Number: 2109671053.

.

Expressions: set # 3

Attributes of the expression 3

Title: Rayuela, Himmel-und-Hoelle: Roman aus dem argentinischen Spanisch

Language: **German Translations**

Form: text

Manifestations

Manifestation 1

Attributes of the Manifestation 1

Title: Rayuela, Himmel-und-Hoelle: Roman aus dem argentinischen Spanisch

Statement of Responsibility: Julio Cortázar

Statement of Responsibility for the Translation: Fritz Rudolf Fries

Place of Publication: Frankfurt

Publisher: Suhrkamp

Date of Publication: 1981

Items from Manifestation 1

Item # 1

Attributes

Item identifiers:

Call Number:. PQ 7797 C7145 R315 1981

Inventory Number:. 2010389543

Conclusions

The FRBR model constitutes an important contribution to library cataloguing because it is an attempt to functionalize the concept of work for depicting relationships among bibliographic items. Also, the FRBR proposal takes into account the entity-relationship database modeling system, which was demonstrated to be an appropriate tool for the development of comprehensive systems of entities from a users' point of view. We celebrate and welcome IFLA's initiative.

However, the proposal can be improved in many ways. I will discuss an evaluation of the FRBR model regarding my experience in processing the work I chose (*Rayuela,* by Julio Cortázar) along with the insights by Edward T. O'Neill (2002), who undertook the same enterprise with the novel *Humphry Clinker*. I will focus on two elements: the makeup of attributes and the functionality of the entity expression. Finally, I will briefly discuss the possible further research to keep on developing the model.

Issue 1: Attributes.

The concept of *attribute* is functional for the user to identify a particular item or set of items that he wants to access. Therefore, the function of identification is essentially related to the definition of each attribute for the entity, whether work, expression, manifestation or item. Each attribute has to be

unequivocally related to the entity that holds it. For users better to identify entities, it is essential not only clearly to define the scope and limits of each entity, but also to avoid multiplication and overlapping of those attributes.

One of my findings when applying the model to actual items is that some attributes do not allow clear differentiation. As far as some attributes are defined similarly, there is repetition in the occurrence for them. For example, there can be duplication of the attribute *title* between entities work and expression. For books, if we consider the first publication ever realized to be the actual work, there would be a multiplication of these attributes in at least three levels of entities: work, one set of expressions (the ones that join the manifestations of the original work), and the manifestations of such expression. We can demonstrate this by looking at the set of manifestations for expression #1, all of which hold the same title. This unnecessary repetition not only is laborious for the cataloguer, but also makes the display confusing and cumbersome for the user.

In a somehow different approach, another problem related to attributes is the lack of some of them. For example, there is no attribute related to the statement of responsibility for translations. I added it when applying the model for the translation by Gregory Rabassa (see Expressions set # 2, English translations). Also, the model does not provide attributes for those cases in which the work has an introductory section attached to it. O'Neill (2002) arranged as new expressions those items that hold supplementing material, such

as maps, chronological tables, and bibliography (p.154). However, FRBR does not deal with such supplements.

Issue 2: The Notion of Expression

O'Neill concludes, after assessing the application of FRBR to *Humphry Clinker*, that replacing expression with additional manifestations attributes will work for his set of items (p. 158). He supports taking out this entity based on the fact that it is hard to identify whether two manifestations have the same content. I agree with his conclusion and add that in reality no cataloguer can test at what point two expressions have the same or similar content that for the sake of effectiveness. A cataloguer has neither the time nor the expertise to evaluate if the content of two manifestations is different (for example, by reading them). However, what all cataloguers can do is to collect and transcribe appropriate attributes for identification and retrieval. Additionally, O'Neill adds that "even if it could easily be determined when the content was identical the result would have an overly fine granularity" (2002, p.158), granularity that goes against a users' call for simplicity and effectiveness. Establishing the entity expression duplicates the level of analysis without bringing further operational benefits to the model.

I should note, however, that with the novel *Rayuela* expressions did not add this kind of complexity. Instead, they allowed me to categorize sets of

manifestations that otherwise would be put together in a confusing set. My hierarchical chain eventually was made up as follows:

Work
- Expression set #1: six manifestations
- Expression set #2: three manifestations
- Expression set #3: one manifestation

Each expression basically represents a different language in which the work is conveyed. Without going further in granularity, we can take advantage of the set of attributes detailed for the entity expression to draw basic categories of manifestations. We have to bear in mind that we are at a very abstract level of categorization when dealing with expressions. The more concrete the level of description, the longer the list of attributes that the item has. The fact that the entity expression has a list of twenty-five attributes does not mean that we need to work with all of them for the particular IE we describe at the moment, but rather only as many of them as are necessary to distinguish expressions from each other.

I believe the entities work and expression act as major categories that allow cataloguers to find relationships among manifestations. The detailed cataloguing description, i.e. the gathering of attributes and setting of occurrences for them, has to be accomplished at the levels of both manifestations and items, where the level of concreteness is higher. Even when manifestation clearly is not

the concrete physical object that we hold at the library (which is the item), this entity joins together those items.

Issue 3: The Scope of *Manifestation's* attributes

Another problematic element to apply in my set of items was the broad number of attributes that the entity *manifestation* has in front of the entity *item.* As we already discussed,[9] in FRBR model the level of abstraction in group-1 entities decreases (i.e., the level of concreteness increases) from the first to the last entity—i.e. from *work,* the more abstract, to *item*, the more concrete. We also analyzed that the relationship among these entities is hierarchical,[10] which means that, for example, an expression is intrinsically and univocally related to only one work. This dependence is strongly related to the level of abstractness-concreteness among entities. In *Rayuela*, the novel I tested, I faced one single work that has three sets of expressions, ten manifestations, and fourteen items. All entities were vertically dependent in the model.

Following this hierarchical structure, as we go down in the vertical array and reach higher levels of concreteness, we face an increasing number of attributes. The number of attributes increases as we go down in the hierarchy because we start identifying particular items, and identification always

[9] See section 2.3.1. Entities
[10] See section 2.3.1. Hierarchical Relationships

encompasses naming attributes. For example, we describe the work *Visions of Cody*[11], by Jack Kerouac naming the attributes *title, form* (novel), *date* (1951-52)[12], *intended audience* (adults) and *context of the work* ("the Beat generation," post war WWII). Following FRBR, our list of attributes for Kerouac's work finishes here. However, when we describe one particular expression for that work, we have to add attributes such as *date* (1972), and *language* (English). Reaching the *manifestation* level, we will find the more detailed number of attributes. We add *edition* ("edition by Penguin Books"), *place of publication* (New York) and *date of publication* (1993). Finally, an *item* of this work will be identified by its *mark or inscriptions* (in the case of the item I am describing, an inscription by a friend who gave it to me as a gift). Therefore, to identify the item I own I have to come up naming all those attributes, following the hierarchical array.

Interestingly, FRBR offers the most extensive list of attributes for the entity *manifestation* (thirty eight attributes compared to twenty five for expressions, twelve for work, and only nine for items). This breadth in the level of detail for *manifestations* makes its list of attributes very extensive and, at the same time, gives a more concrete level of ontology to this entity, level of concreteness that in reality belongs to the entity *item*. Going back to the novel I

[11] I will describe this novel based on a particular item that I owe.
[12] The novel was written by Kerouac in 1951-52 but not published until 1972.

tested, when applying FRBR model I had, in expression set # 1, seven manifestations (manifestations # 1 through # 7) that are in fact printings of the same expression of the work, made by the same publisher, in the same place, etc. The only datum that varies is the *date*[13] (1963, 1968, 1969, 1974, 1975, 1976, 1977). Even more, when we see the items we realize that they have the same pagination, format, and typography. In conclusion, these seven items differ only by the date of publishing. But FRBR allows the cataloguer to detail the date in the most concrete level within the attributes for a *manifestation,* not for an *item.* According to my experience, FRBR model would be more consistent if it had attribute *date* among the *item's* attributes—not in the *manifestation* list.

Now I will display the same set of items but placing the attribute *date* in the *item's* list:

EXPRESSIONS: SET # 1: ORIGINAL WORK
Attributes of the expression
Title: Rayuela
Language: **Spanish Versions**
Form: Written word

Manifestations
Manifestation 1
Attributes of the Manifestation 1

[13] We should remember that the concept of "edicion" in Spanish is the same as in "printing" in English. Therefore, the mention of 1st, 9th, 11th, 17th, 18th, 19th and 21st editions means, in fact, 1st edition and 9th, 11th, 17th, 18th, 19th and 21st *printings* of the 1st edition (see section 2.3.3.Relationships, second paragraph)

Title: Rayuela
Statement of Responsibility: Julio Cortázar
Place of Publication: Buenos Aires
Publisher: Sudamericana
Date of Publication: 1963

Items from Manifestation 1

Item # 1

Attributes

No. of Edition/Printing: 1st
Date of Manufacturing/Printing: 1963
Item identifiers:
Call Number: PQ 7797 C7145 R3 1963

Item # 2

Attributes

No. of Edition/Printing: 9th
Date of Manufacturing/Printing: 1968
Item identifiers:
Call Number: PQ 7797 C7145 R3 1968 LAC
Inventory Number:. 3002167624
Provenance of the item: Presented by Beverly Gibbs

Item # 3

Attributes

No. of Edition/Printing: 11th
Date of Manufacturing/Printing: 1969
Item identifiers:
Call Number: PQ 7797 C7145 R3 1969 LAC
Inventory Number: 2114543110
Provenance of the item: Presented by Beverly Gibbs

Item # 4

Attributes

No. of Edition/Printing: 17th
Date of Manufacturing/Printing: 1974
Item identifiers:
Call Number: PQ 7797 C7145 R3 1974 LAC
Inventory Number: 2013274821

Item # 5.1
Attributes
No. of Edition/Printing: 18th
Date of Manufacturing/Printing: 1975
Item identifiers:
Call Number: PQ 7797 C7145 R3 1975 LAC
Inventory Number: 2013274832

Item # 5.2
Attributes
No. of Edition/Printing: 18th
Date of Manufacturing/Printing: 1975
Item identifiers:
Call Number: PQ 7797 C7145 R3 1975 LAC COPY 2
Inventory Number: 3020512746

Item # 6
Attributes
No. of Edition/Printing: 19th
Date of Manufacturing/Printing: 1976
Date: 1976
Item identifiers:
Call Number: PQ 7797 C7145 R3 1976 LAC
Inventory Number: 2013274843

Item # 7
Attributes
No. of Edition/Printing: 21th
Date of Manufacturing/Printing: 1977
Item identifiers:
Call Number: PQ 7797 C7145 R3 1977B LAC COPY 2
Inventory Number: 2102012591

In the first set of records that I made up (section 2.1), starting from manifestation # 8, several attributes change such as *place of publication* and *publisher.* Therefore, they belong to a new set of manifestations.

As the reader might notice, item #5 in this new display has two sub-items levels. This division has to be done because the set holds two items of year 1975's printing. I think that it is better to split up at the level of item than to broaden at attributes at the level of manifestations.

The Manuscript

A special note has to be given to the fact that the Benson Latin American Collection (BLAC) owns the manuscript of the work, *Rayuela*. I have not included it in my analysis since the categorization of manuscripts is not clear in FRBR. However, I examined the manuscript and found that its content is different from that which the novel had after being published. The pages of the manuscript show how the author wrote and re-wrote each sentence of the novel, striking through the text and writing different versions of the same sentences. Therefore, in the case of this manuscript we face different wordings for the same text. The *content*, which determines the identity of the *work,* is controversial, and discussing it would require an analysis that goes beyond the scope of this report. Even so, I would like to point out that the manuscript of *Rayuela* could be related to the final version of the novel (the *work* that I tested), through a *work-to-work* relationship.[14] The *work-to-work* horizontal relationship includes categories such as *successor, supplement* or *complement* (FRBR, 1998, p. 65).

The manuscript of *Rayuela* could be included in a new category such as *predecessor,* which would allow showing the natural relationship between the works. The building of these relationships would also offer a functional approach to Svenonius' concept of *superwork* because such connections provide a way of relating works with *similar,* though neither *equal* nor *essential* content[15].

Summary

The FRBR model has its roots in an extensive research in cataloguing. The notion of *work,* a crucial concept in the model, has been explored both ontologically and functionally since the 1950s by a wide range of experts. FRBR takes the findings of this exploration and adds an operational approach brought from the database management science. The *entity-relationship database system modeling* provides FRBR with an operative scheme which fixes concepts related to the description of informational items.

The result is a flexible, user-centered model for creating surrogate records in cataloguing. However, the model has to be explored thoroughly. Applying FRBR in a set of items related to the same work is an effective way of testing how the system does well, and which aspects should be improved.

---

[14] See section 2.3.3.1.Other Relationships
[15] See sectioin 1.2.2. Superwork

Through this report, I have contributed to the exploration of FRBR by modeling a text-based IE in a non-English language. I close my research looking forward to additional experiences in the application of the model in our worldwide cataloguing community.

References

Bopp, R. & Smith, L. (2001). *Reference and information services: An introduction.* Eglewood, CO: Libraries Unlimited.

IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998). *Functional requirements for bibliographic records: Final report.* International Federation of Library Associations. Retrieved February 2, 2003, from the International Federation of Library Associations Web site: http://www.ifla.org/VII/s13/frbr/frbr.pdf

Leazer, G. & Smiraglia, P. (1999). Bibliographic families in the library catalog: A qualitative analysis and grounded theory. In *LRTS, 43*(4), 191-213.

Lubetzky, S. (1969). *Final report: Phase I, descriptive cataloging.* Los Angeles, CA: University of California, Institute of Library Research.

Mayhew, D. (1999). *The usability engineering lifecycle.* San Francisco, CA: Morgan Kaufmann.

Miksa, F. (2000). *A general outline of information entity access control*. Retrieved October 4, 2002, from University of Texas at Austin, LIS 384K.8 Blackboard Site

O'Neill, E. (2002). FRBR: Functional requirements for bibliographic records: Application of the entity-relationship model to *Humphry Clinker.* In *LRTS, 46*(4), 150-159.

Rob, P. & Coronell, C. (2002). *Database systems: Design, implementation and management* (5th ed). Boston, MA: Thomson Learning.

Rosenfeld, L. & Morville, P. (2002). *Information architecture for the world wide web* (2nd ed.). Cambridge: O'Reilly.

Smiraglia, R. (2001). *The nature of "a work": Implications for the organization of knowledge*. Maryland: The Scarecrow Press.

Svenonius, E. (2000). *The intellectual foundation of information organization*. Cambridge, MA: MIT Press.

VITA

Norma Estela Palomino was born in Haedo, Buenos Aires, Argentina on May 25, 1970, the daughter of Ofelia Mabel Pisano and Rodolfo Javier Palomino. After completing her work at Colegio Nacional Esteban Echeverria de Ramos Mejia, Buenos Aires, in 1987, she entered Universidad de Moron in Moron, Buenos Aires. She received the degree of Profesora de Filosofia in April 1993. During the following years, she held several library positions in Argentinean Universities. In August 2001, she entered The School of Information at the University of Texas.

Permanent Address: Gurruchaga 556, 1ro. "J"
1414 – Capital Federal
Argentina

This report was typed by the author